

REAL TIME MONITORING OF SURVEILLANCE CCTV

NADHIRAH BT ALI

Thesis submitted in fulfilment of the requirements
for the award of the degree of
Bachelor of Electric Electronics Engineering (Electronics)

Faculty of Electrical and Electronic Engineering
UNIVERSITI MALAYSIA PAHANG

NOVEMBER 2010

SUPERVISOR'S DECLARATION

I hereby declare that I have checked this project and in my opinion, this project is adequate in terms of scope and quality for the award of the degree of Bachelor of Electrical & Electronics Engineering (Electronics).

Signature

Name of Supervisor: DR. KAMARUL HAWARI BIN GHAZALI

Date: 29 NOVEMBER 2010

STUDENT'S DECLARATION

I hereby declare that the work in this project is my own except for quotations and summaries which have been duly acknowledged. The project has not been accepted for any degree and is not concurrently submitted for award of other degree.

Signature

Name: NADHIRAH BT ALI

Date: 29 NOVEMBER 2010

ACKNOWLEDGEMENTS

I take this opportunity to thanks my advisor, Dr Kamarul Hawari Bin Ghazali for his guidance throughout my thesis. It was his word of encouragement that help me to perceive my research goal. I thank the entire course mate for their time and patients in reviewing my thesis.

For the unending love and understanding from my parent who motivated me to accomplish this through, and I am very grateful to my sibling who give me the courage and strength.

I would like to take this opportunity to thanks to my friends, who give me support and help.

ABSTRACT

Traditionally the surveillance monitoring system is done in the larger room and by a mount of manpower. But nowadays, monitoring surveillance system can be done through online network. This type of monitoring is more time consuming and can reduce the manpower. Moreover, it gives the user flexibility to monitor their properties where ever they want as long as they have the internet network. Either than that, this project also manage to detect the movement in the video, this will be a greater help to the user to observe their properties. In the end, this project is not only can reduce the cost for monitoring but also give advantages to the user.

ABSTRAK

Secara tradisional sistem pemantauan kamera litar tertutup dilakukan di sebuah bilik kawalan yang besar bersama dengan jumlah tenaga kerja yang ramai. Tetapi pada zaman yang lebih maju ini, sistem ini dapat dilakukan melalui rangkaian internet. Dengan cara ini ia dapat mengurangkan masa dan tenaga kerja bagi satu – satu pemantauan. Selain daripada itu, sistem ini memberikan fleksibiliti kepada pengguna untuk memantau harta mereka di mana mereka berada dengan syarat mereka mempunyai rangkaian internet. Selain dari itu, projek ini juga boleh mengesan gerakan dalam video, dan ini akan menjadi bantuan kepada pengguna untuk memantau harta mereka dengan lebih mudah. Keimpulannya, projek ini bukan sahaja dapat mengurangkan kos untuk pemantauan tetapi juga memberikan kelebihan kepada pengguna.

TABLE OF CONTENTS

	Page
SUPERVISOR’S DECLARATION	ii
STUDENT’S DECLARATION	iii
ACKNOWLEDGEMENTS	v
ABSTRACT	vi
ABSTRAK	vii
TABLE OF CONTENTS	viii
LIST OF FIGURES	x
CHAPTER 1 INTRODUCTION	
1.1 Introduction	1
1.2 Problem Statement	3
1.3 Project Objective	3
1.4 Project Scope	3
CHAPTER 2 LITERATURE REVIEW	
2.1 Introduction	4
2.2 Motion Detection	4
2.3 Remote Monitoring	7
2.4 Real Time Detection	8
2.5 System Network	11

CHAPTER 3 METHODOLOGY

3.1	Introduction	13
3.2	Methodology	13
	3.2.1 Setting the IP Camera	13
	3.2.2 Image Processing	14
	3.2.3 Graphical User Interface	17
	3.2.4 Website	17
3.3	Technique	18
	3.3.1 IP Address	18
	3.3.2 Image Processing	18
	3.3.3 Graphical User Interface	21

CHAPTER 4 RESULTS AND DISCUSSION

4.1	Introduction	23
4.2	Offline Detection	23
4.3	Real – Time Monitoring	24
4.4	Discussion	28

CHAPTER 5 CONCLUSION AND RECOMMENDATIONS

5.1	Conclusions	29
5.2	Recommendations	30

REFERENCES	31
-------------------	----

APPENDICES	33
-------------------	----

A	Gantt Chart	33
B	Source Code	34
C	Result Output	37

LIST OF FIGURES

Figure No.	Title	Page
1	Block diagram of algorithm	8
2	System architecture	11
3	System architecture	12
4	RGB format of the video	23
5	Black and white format of the video	24
6	Rectangle around the object	24
7	Layout of the GUI	25
8	Snapshot figure	26
9	Example of image save in the memory	26
10	Indicator the video is created	27
11	File save in the memory	37

CHAPTER 1

INTRODUCTION

1.1 PROJECT BACKGROUND.

CCTV is the acronym of Closed-circuit television (CCTV). Surveillance CCTV is one of the most important evidence when deals with wrong doings. A video surveillance system covering a large office building or a busy airport can apply hundreds and even thousands of cameras. To avoid communication bottlenecks, the acquired video is often compressed by a local processor within the camera, or at a nearby video server. The compressed video is then transmitted to a central facility for storage and display.

Based on the current technologies, with the set of personal computer (PC) and the internet connection either wire or wireless, the monitoring can be done. With this, the user may monitor the video wherever they want, and the random video playback functions can be provided. With these flexibilities, it gives more advantage to the user to monitor and ensure the safety place they want. It is also may increase the safety of the user properties; this is because there is image processing technique apply in the system.

Image processing apply in this system is to detect object movement. By using this technique it will tell the user there is movement in that particular frame. This technique will highlight movement, thus will alert the user. To apply this technique some analysis must be done. In this project, the analysis is done by using MATLAB software. MATLAB is software that widely used in the computer analysis circle. In the MATLAB there is image processing toolbox. This toolbox will guide in the image processing technique and the details about the concept apply.

The Graphical User Interface (GUI) is built in the MATLAB. This GUI is build for user friendly purpose. The built of the GUI will guide the user to monitor their CCTV. It will make programs easier to use by providing them with a consistent appearance and with intuitive controls like pushbuttons, list boxes, sliders, menus, and so forth. All of these elements are complete in MATLAB GUI builder.

1.2 PROBLEM STATEMENT

Even though, the monitoring can be done remotely, with some help in image processing will increase the efficiency of the system. In this project, image processing is applied to let the user know there is object movement. This will reduce time consuming of the user. This will make the monitoring is a lot mo easier.

This project is a web base project, but the analysis is done in the MATLAB format, thus there is conflict in the language used. There for, studies in web language are necessary in order to called back the MATLAB function in web language.

1.3 PROJECT OBJECTIVE

The objective needs to archive to complete this project are:-

1. Obtain the IP address of the IP base CCTV.
2. Build the GUI for user usage for real – time monitoring.

1.4 PROJECT SCOPE

These scopes are determined in order to complete this project:-

1. Apply the image processing technique by using MATLAB 2010a.
2. Design the GUI for real time monitoring
3. Studies of web language for web site design.

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

The purpose of this chapter is to discuss of literature review about image processing and remote monitoring using web base.

2.2 MOTION DETECTION

Motion detection is the most commonly technique use nowadays. This is because their usage in many areas like video surveillance systems, traffic monitoring, gesture recognition, advanced user interfaces, sport games players tracking etc. All these application is to identify the object of interest. There are a number of proposed methods in the literature for moving object and all of them can be divided into three classes: temporal difference, optical flow and background subtraction [1].

The background subtraction methods are the most commonly used with the static cameras because of its high performance and low memory requirements [1]. Background subtraction is a method typically used to segment moving regions in video sequences by comparing each new frame to a model of the scene background. It has been used successfully for indoor and outdoor applications [2]. This method need the background and the camera remain static even though the background color intensity may change gradually base on the current situation. Thus it is necessary to compare pixel of the current frame background with the corresponding background.

This is the drawback of the background subtraction [1]. This is because due to the background scenes often change, such a car running into or out of the background, someone brings things into or out of the background or even the wavering leaves. For effectiveness and accuracy in this method, the initialization and the background update is important [3].

Thus, based on the idea of updating the background, the adaptive background method is proposed. Adaptive background subtraction (ABS) is a fundamental step for foreground object detection in many real-time video surveillance systems. In many ABS methods, a pixel-based statistical model is used for the background and each pixel is updated online to adapt to various background changes. But, by using this method it needs a lot of computational and memory consumption required. Thus in the paper [4] propose the usage of K-Means clustering method to adaptive background model which is constructed by the Mixture of Gaussians (MOG) for improving the computational time.

As for the busy situation, background subtraction is irrelevant to be applied for motion detection, thus optical flow method is introduced. Optical flow is an approximation of the local image motion and specifies how much each image pixel moves between adjacent images. It can achieve success of motion detection in the presence of camera motion or background changing. According to the smoothness constraint, the corresponding points in the two successive frames should not move more than a few pixels. For an uncertain environment, this means that the camera motion or background changing should be relatively small. The method based on optical flow is complex, but it can detect the motion accurately even without knowing the background [5].

Most optical flow techniques are either gradient based methods [6] or region matching based methods [7]. Gradient based methods have been preferred due to speed and performance considerations. These methods analyse the change in intensity and gradient (using partial spatial and temporal derivatives) to determine the optical flow and it also yields smoother and more natural motion fields [7,8]. The popular optical flow technique; Horn and Schunck are a combination of the gradient constraint with a global smoothness term [7] to be smooth in the edge region, the speed of which always sharp changes, so the edge shape of the object easily distorts to constrain the velocity field [7,9]. As for the region matching it relies on determining the correspondence between the two images, by matching 'blocks' of one image to 'blocks' of the other [8].

By using optical flow, detection method of a moving object by mapping, which converts the motion of a stationary environment object into a linear signal trajectory. The one-dimensional optical flow is calculated by using pixels, which belong to the moving object, to eliminate the apparent motion of the stationary environment object [10]. By using optical flow also can detect a multiple motion in a frame because it is calculated based on the velocity of the particular object [8]. Based on this characteristic, it can be used to extract the candidate region based on matching optical flow values [11].

As for temporal difference, it computes the difference between two or three consecutive frames. It is good at adapting to the dynamic environments, but generally poor at extracting enough relevant feature pixels, which resulting holes are generated in the moving object [12] and cannot detect the entire shape of a moving object with uniform intensity [13].

2.3 REMOTELY MONITORING

Monitoring remotely is analogy from the conventional monitoring technique. In the control room, there are many monitor display the condition of the respective area, same goes for online monitoring. For online monitoring there is setting that can be change either to watch single video or multiple video. In the digital CCTV monitoring also can have apply processing in there. Thus, we can detect the movement, either it is normal or abnormal. There is study in this area says that, the electronic engineers have an understanding of field theory and of flow dynamics which may provide insight into characteristics of crowd behavior, and can also provide the expertise to suggest solutions to crowd monitoring and control based on technological developments in image processing and image understanding [14].

IP-based camera used in a remote monitoring station for intelligent image processing. In order to provide enhanced security, events need to be integrated with a command and control system capable of effectively responding to hundreds of events per day in a busy, critical infrastructure facility [8]. Other than IP-based camera, pan, tilt, and zoom (PTZ) camera also can be used in remote monitoring. The PTZ itself may refer to the feature for the security camera. It also can be used to describe an entire category of cameras where a combination of sound and/or motion and/or change in heat signature may enable the camera to activate, focus and track suspected changes in the video field.

2.4 REAL TIME DETECTION

The real-time abnormal motion detection scheme uses algorithm the macro block motion vectors that are generated anyway as part of standard video compression methods. Motion features are derived from the motion vectors. Normal activity is characterized by the joint statistical distribution of the motion features, estimated during a training phase at the inspected site. During online operation, improbable-motion feature values indicate abnormal motion. Relying on motion vectors rather than on pixel data reduces the input data rate by about two orders of magnitude, and allows real-time operation on limited computational platforms [20].

In paper [1] purpose algorithm base on this block diagram.

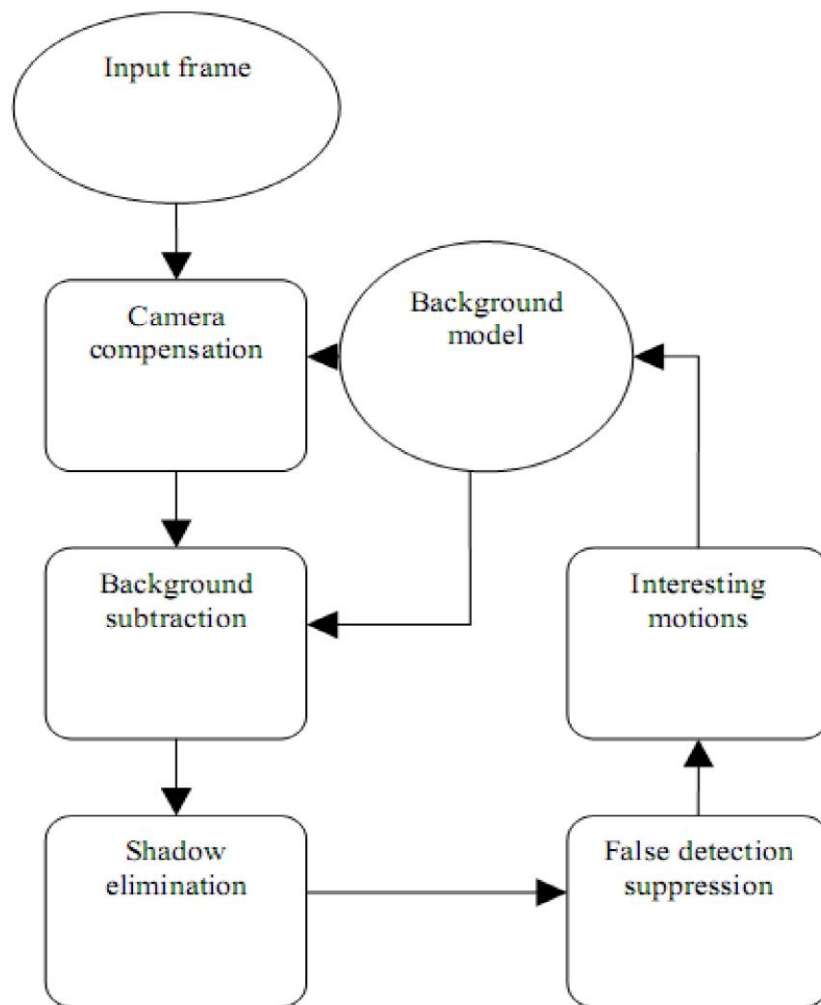


Figure 1: block diagram of algorithm [1]

The camera motion of the input frame is compensated and consecutive temporal difference is performed to extract moving areas from the image. The moving areas include the target areas (moving objects we are interested in) and some uninteresting motion areas (the moving background). Post treatments are used to optimize the detection by filling the small holes in the detected objects and removing uninteresting motions. Since shadows won't have large change between two consecutive frames and little change of shadows to be detected can be removed by the post treatments, shadows have little effect on the accuracy of detection. Thus, they are not handled here to improve the efficiency of this method. To perform this method in real-time and with high accuracy, we design every block carefully. Since the consecutive temporal difference approach requires no background model and little memory, this approach is quite efficient and accurate in that it has a low computational cost and it adapts quickly to the changes of the background.

First, the input frame is compensated by using a camera motion compensation algorithm, which takes some spots in the frame as basic pixels and estimates the motion of the camera with a square neighborhood matching method. The whole input frame is then adjusted to make up the motion of the camera. The square neighborhood matching method will be fully described in the following subsection. It shows how it records the deviation of the image.

Second, an improved consecutive temporal difference approach is used to quickly obtain moving areas of the input frame. This approach makes use of three consecutive frames. The three frames are divided into two groups. The first group includes the two previous frames (the two consecutive frames that go before the input frame), while the second group includes the input frame and the frame before it. By subtracting the two groups separately, we get two results of different areas from the two subtractions. Since the intersection of the two results is just the very part of the moving area in the frame previous to the input frame, we can obtain the moving areas by subtracting the second results by the intersection of the two difference frames

Third, post treatments are used in this phase to fill the small holes in the detected objects and to remove uninteresting motions from the moving areas obtained in the second phase. Since shadows are not handled in this method, post treatments deal mostly with the two problems mentioned above. The math morphology techniques are used.

The Algorithm Based Object Recognition and Tracking (ABORAT) system presented in this paper is a vision-based intelligent surveillance system, capable of analyzing video streams. These streams are continuously monitored in specific situations for several days (even weeks), learning to characterize the actions taking place there. This system also infers whether events present a threat that should be signaled to a human operator. The concept of the ABORAT system is to apply intelligent vision algorithms on images acquired at the system's edge (the camera), thus reducing the workload of the processor at the monitoring station and the network traffic for transferring high resolution images to the monitoring station [15].

A layered architecture for real-time surveillance systems which for each layer includes objects that model the "real world" at a specific abstraction level, from raw data up to domain concepts. Each layer performs abstractions on perceptions coming from the lower layer and formulates timed hypotheses about domain objects. The failure of a hypothesis causes a perception to flow up-stream. In turn, hypotheses flow down-stream, so that their verification is delegated to the lower layers. The proposed architectural patterns have been reified in a Java framework, which has been used in an experimental multi-camera tracking system [16].

For multiple object detection to extract and do well reveal the foreground object, the process of object detection in the proposed algorithm is based on the background subtraction and the double-difference image. Typically, background subtraction is the first step in automated visual surveillance applications and a method used to segment moving regions in video sequences taken by a camera by comparing each new frame to the scene background model. We first create a clean background model which is similar to that taken by an adaptive background modeling algorithm based on a Gaussian-mixture model . This model

can reduce the noise generated from a camera itself or from the lamplight twinkling [17].

2.5 SYSTEM NETWORK

The system architecture of the present networked visual monitoring system is shown in Fig. 8. There are types of two cameras in the monitored site. One is the global camera, which captures the global view of interest. The other are tracking cameras, controlled by VMS server to track the intruders. In order to keep the recorded images safely, the captured images are stored in a VMS server as well as VMS DB server. In the case users request real time image monitoring, the system connects to the VMS Server. On the other hand, the system connects to the VMS DB server if users query stored images [18].

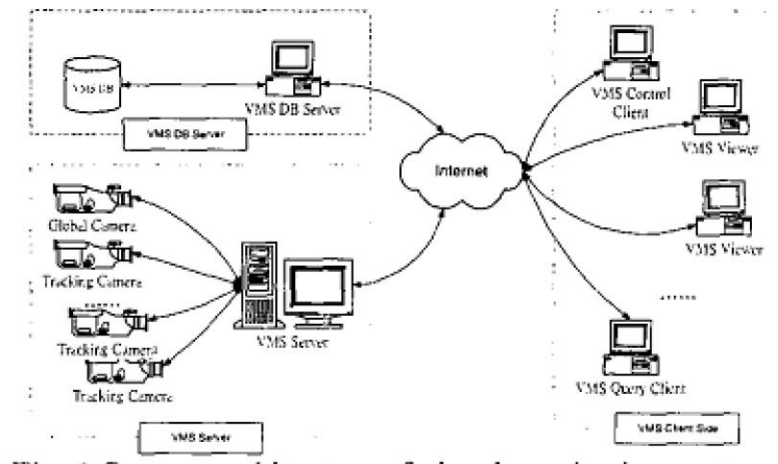


Figure 2: system architecture [18].

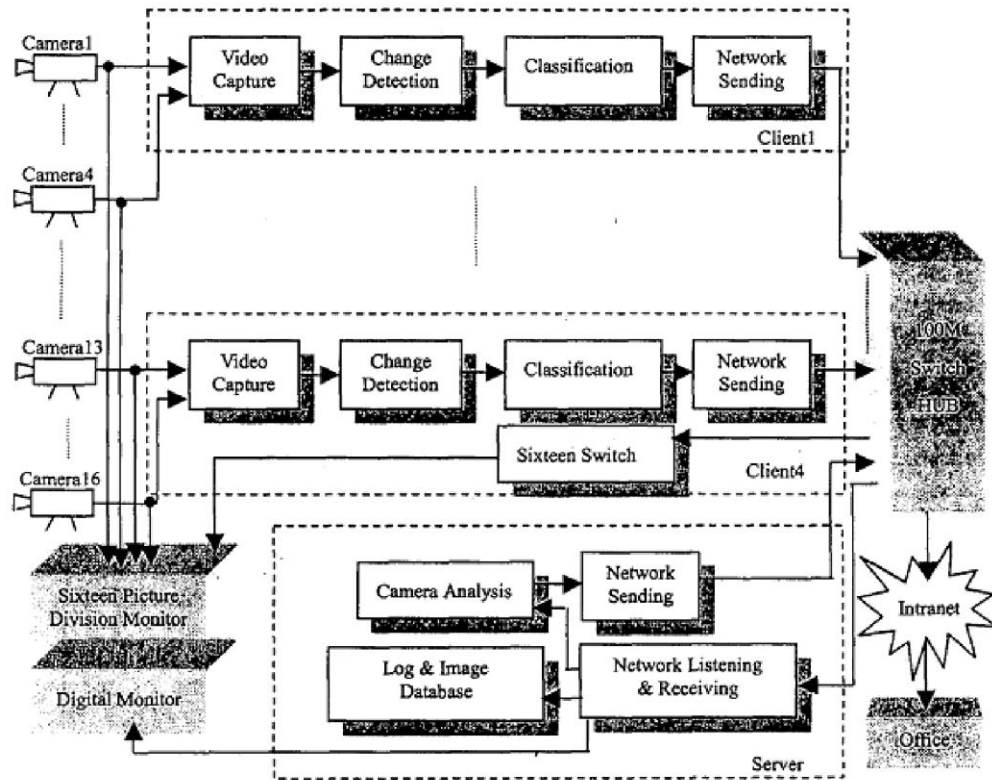


Figure 3: system architecture for [19]

A network-based visual intelligent surveillance system is mainly composed of cameras, client computers, a server computer and a HUB. The clients include video capture module, change detection module, classification module and network sending module. The server includes network listening & receiving module, camera analysis module, a digital monitor and log & image database. The architecture of a surveillance system with sixteen cameras is shown in Fig.9. The video capture Module on each client can capture video sequence from four cameras. So there are four clients in all. All sixteen cameras are connected to the corresponding video capture card on the client computers on the one hand, and connected to an analog sixteen-picture-division monitor on the other hand. On the natural situation, sixteen pictures come from sixteen cameras are simultaneously shown on the analog division monitor [19].

CHAPTER 3

METHODOLOGY

3.1 INTRODUCTION

In this chapter will describe the details steps taken to detect motion movement analysis. The details also will cover for real time monitoring and how web base is build.

3.2 METHODOLOGY

3.2.1 Setting the IP camera.

The first step is to setup the IP camera at the desire place. After place the IP camera is to obtain the IP address of the IP camera. This is importance since to connect the IP camera with the web is the IP address. IP address is a numerical label that is assigned to devices participating in a computer network that uses the Internet Protocol for communication.

3.2.2 Image processing.

Mean filtering is a simple, intuitive and easy to implement method of smoothing images, such as to reducing the amount of intensity variation between one pixel and the next. It is often used to reduce noise in images. It is simply to replace each pixel value in an image with the mean ('average') value of its neighbors, including itself. This has the effect of eliminating pixel values which are unrepresentative of their surroundings. Mean filtering is usually thought of as a convolution filter. Like other convolutions it is based around a kernel, which represents the shape and size of the neighborhood to be sampled when calculating the mean. Often a 3×3 square kernel is used.

The Gaussian smoothing operator is a 2-D convolution operator that is used to 'blur' images and remove detail and noise. In this sense it is similar to the mean filter, but it uses a different kernel that represents the shape of a Gaussian ('bell-shaped') hump. This kernel has some special properties which are A box is scanned over the whole image and the pixel value calculated from the standard deviation of Gaussian is stored in the central element.

The median filter is normally used to reduce noise in an image, somewhat like the mean filter. However, it often does a better job than the mean filter of preserving useful detail in the image. Noise is removed by calculating the median from all its box elements and stores the value to the central element.

Image segmentation refers to the process of dividing image into regions with characteristics, extracting the targets of interest and deleting the useless part. It is one of the most basis and important image processing issue for pattern recognition and low-level computer vision [20].

Histogram based segmentation is one of the easiest way to do image segmentation. Histogram is computed for all the image pixels. The peaks in histogram are produced by the intensity values that are produced after applying the threshold and clustering. The pixel value is used to locate the regions in the image. Based on histogram values and threshold we can classify the low intensity values as object and the high values are background image (most of the cases).

Single gaussian background model is used to separate the background and foreground objects. It is a statically method of separation. In this a set of frames (previous frames) are taken and the calculation is done for separation. The separation is performed by calculating the mean and variance at each pixel position.

Frame difference calculates the difference between 2 frames at every pixel position and store the absolute difference. It is used to visualize the moving objects in a sequence of frames. It takes very less memory for performing the calculation.

Feature Extraction plays a major role to detect the moving objects in sequence of frames. When the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant (much data, but not much information) then the input data will be transformed into a reduced representation set of features. Every object has a specific feature like color or shape.

The most essential feature of image is the image edge. The edge detection is widely applied in the image recognition, image division, image enhancement and image compress and so on, and it is always the focus in the area of digital image processing. Edges are formed where there is a sharp change in the intensity of images. If there is an object, the pixel positions of the object boundary are stored and in the next sequence of frames this position is verified. Corner based algorithm uses the pixel position of edges for defining and tracking of objects.